

Market Segmentation Analysis of Electric Vehicles Market in India

Firdaus Mazumdar

20/09/2024

Case Study:

<https://github.com/ixfirdaus22/FeynnLabsInternship2024/tree/main/Market%20Segmentation>

1. Problem Statement

The tasks involve the working for an Electric Vehicle (EV) startup that is in the early stages of deciding which customer and vehicle segments to target for development. To analyze the Indian EV market using segmentation analysis and provide a strategic recommendation on the most suitable market segment to enter. The goal is to identify which geographic locations, demographic profiles, psychographic traits, and behavioral patterns are most likely to adopt electric vehicles in the early stages of market development.

2. Fermi Estimation

1. **Population of India:** Approx. 1.4 billion people.
2. **Urban Population:** About 35% of the population lives in urban areas.

$$1.4 \text{ billion} \times 0.35 = 490 \text{ million urban population.}$$

3. **Population of Tier-1 Cities:** Major Tier-1 cities like Mumbai, Delhi, Bengaluru, Hyderabad, Chennai, Pune, etc., make up around 15% of India's population.

$$1.4 \text{ billion} \times 0.15 = 210 \text{ million population in Tier-1 cities.}$$

4. **Addressable Market (Target Segment):** Let's assume that 20% of Tier-1 city residents are potential EV buyers based on income, environmental concerns, and tech affinity.

$$210 \text{ million} \times 0.20 = 42 \text{ million potential EV buyers.}$$

5. **Early Adopters:** According to the Innovation Adoption Life Cycle, early adopters make up roughly 16% of the total market.

$$42 \text{ million} \times 0.16 = 6.72 \text{ million early adopters of EVs.}$$

6. Estimated Market Size:

- **Two-wheelers:** Let's assume 40% of the early adopters are interested in two-wheelers.

$6.72 \text{ million} \times 0.40 = 2.69 \text{ million potential two-wheeler EV buyers.}$

- **Four-wheelers:** Assume 30% of early adopters are interested in four-wheelers.

$6.72 \text{ million} \times 0.30 = 2.02 \text{ million potential four-wheeler EV buyers.}$

- **Commercial EVs (B2B):** Let's say 30% of early adopters are interested in commercial EVs (delivery vehicles, etc.).

$6.72 \text{ million} \times 0.30 = 2.02 \text{ million potential B2B EV buyers.}$

This rough Fermi estimation helps approximate the potential market size for different segments within the early adoption phase, allowing for more targeted marketing strategies.

3. Data Collection

Detailed India EV Market Data 2001 – 2024: <https://www.kaggle.com/datasets/srinrealyf/india-ev-market-data>

This dataset provides a detailed overview of the electric vehicle (EV) market in India from 2001 to 2024. It includes monthly sales data, sales data categorized by manufacturer, and vehicle class-wise sales data for different manufacturers. This rich dataset is ideal for market analysis, trend forecasting, and research on the adoption and growth of electric vehicles in India.

- **EV Maker by Place:** List of popular EV Makers and their location of Manufacturing Plant
- **Operational PC:** Total Operational Public Charging Station for EV available in each state
- **Vehicle Class:** Total vehicles (includes electric and all other fuels) registered (manufactured) by category from 2001 - Aug 2024
- **ev_cat_01-24:** Total electric vehicles manufactured from 2001 - Aug 2024 and vehicle category
- **ev_sales_by_makers_and_cat_15-24:** Total electric vehicles manufactured by makers from 2015 - Aug 2024 with the vehicle class

4. Data Preprocessing

a. Handling Missing Data: Techniques like imputation (mean/mode/median) for missing values or remove rows/columns if the missing data percentage is significant.

b. Data Cleaning: consistent formats for dates (e.g., YYYY-MM), numerical data (e.g., sales figures, charging station counts), and categorical data (e.g., manufacturer names).

- **Outlier Detection and Handling:** Identify any anomalies (e.g., extreme spikes in sales) and assess whether they are valid data points or errors.

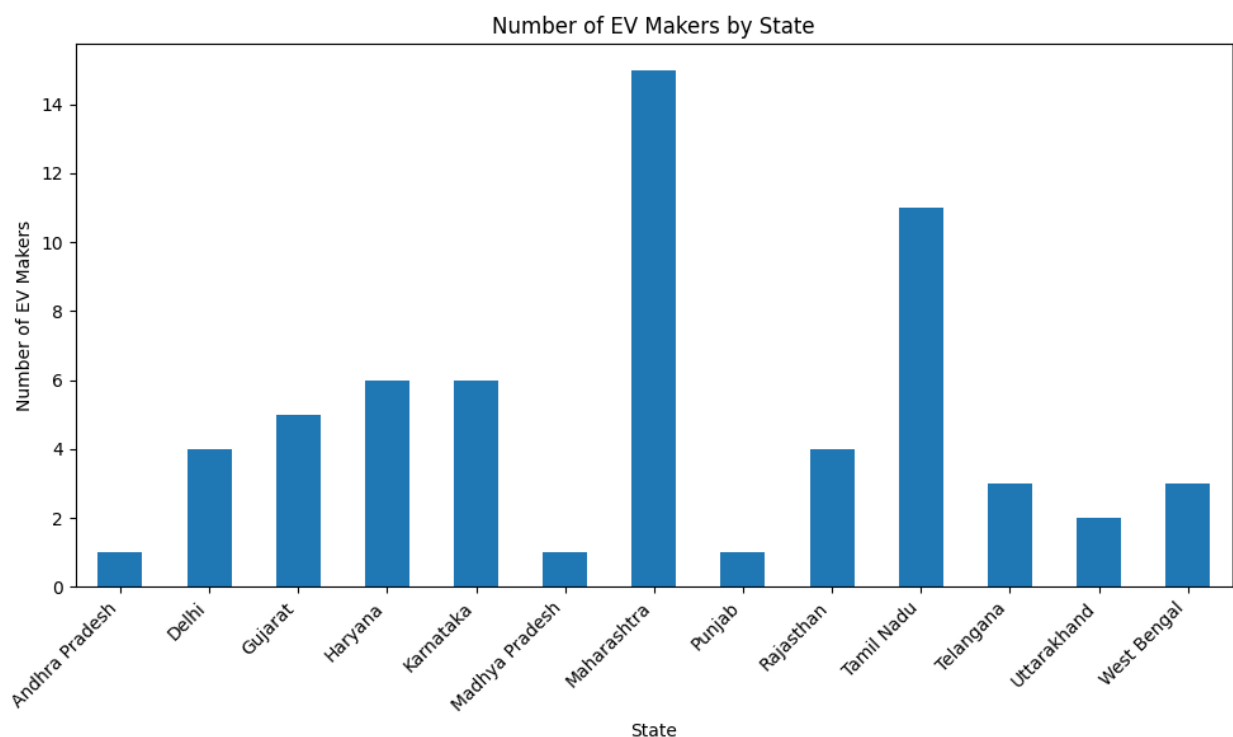
c. Feature Engineering: Add useful features such as "EV Market Share" (EV sales / total vehicle sales), "Charging Station Density" (charging stations per capita in each region), or "Growth Rate" (percentage increase in sales year-on-year).

d. Data Normalization/Scaling: Normalize numerical data to ensure that all features contribute equally to model training, especially when using algorithms like K-Means clustering or logistic regression.

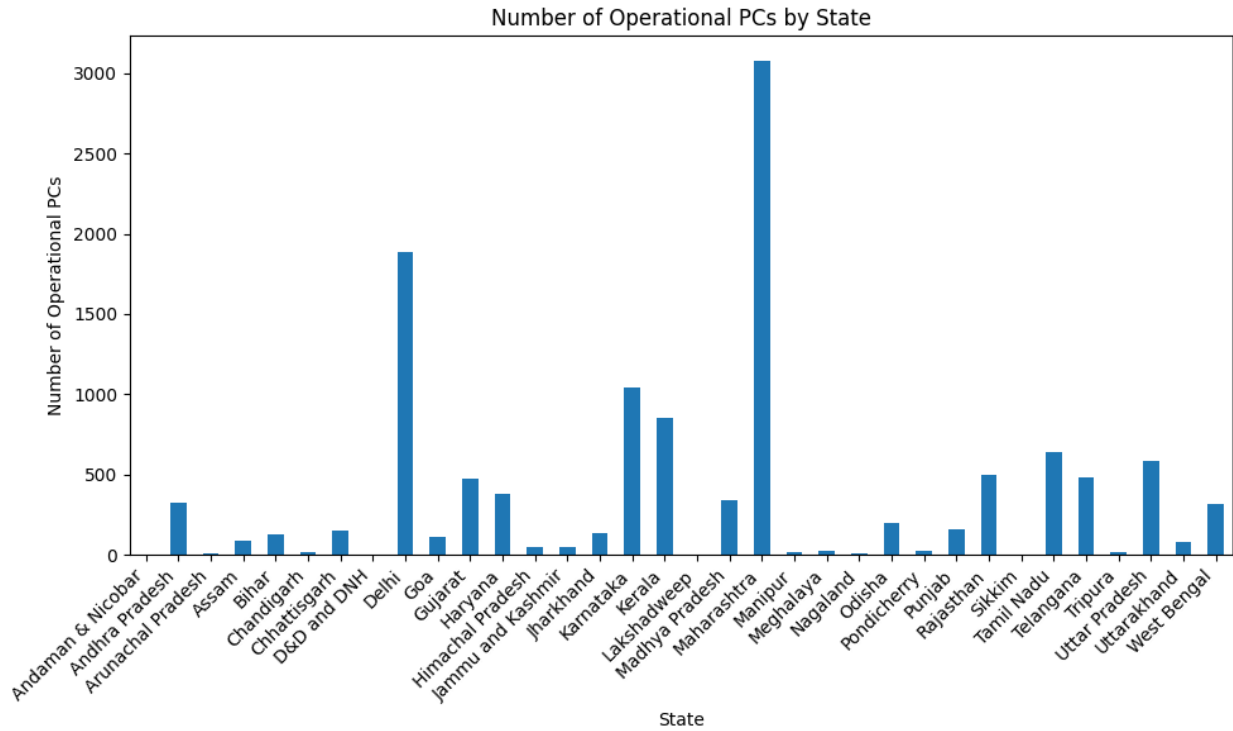
e. Data Splitting: Split the dataset into training and test sets, with around 70-80% data for training and the rest for testing/validation.

5. Exploratory Data Analysis (EDA)

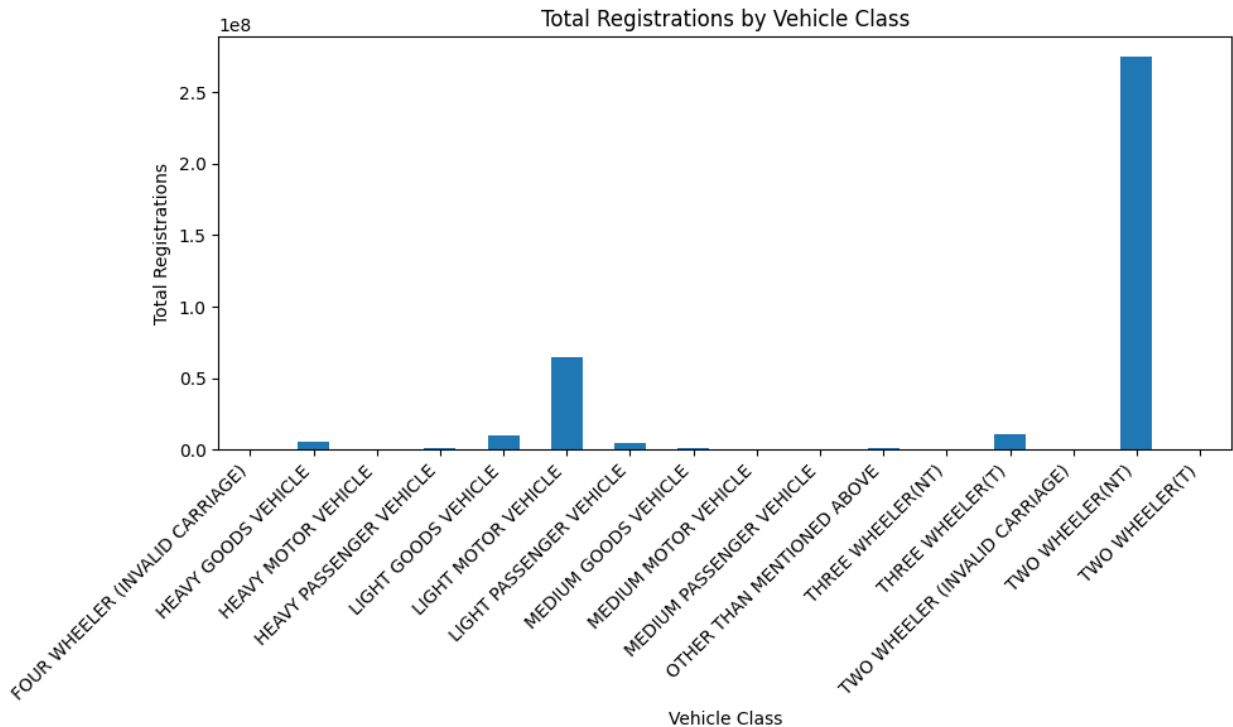
EDA helps uncover patterns, trends, and relationships in the dataset.



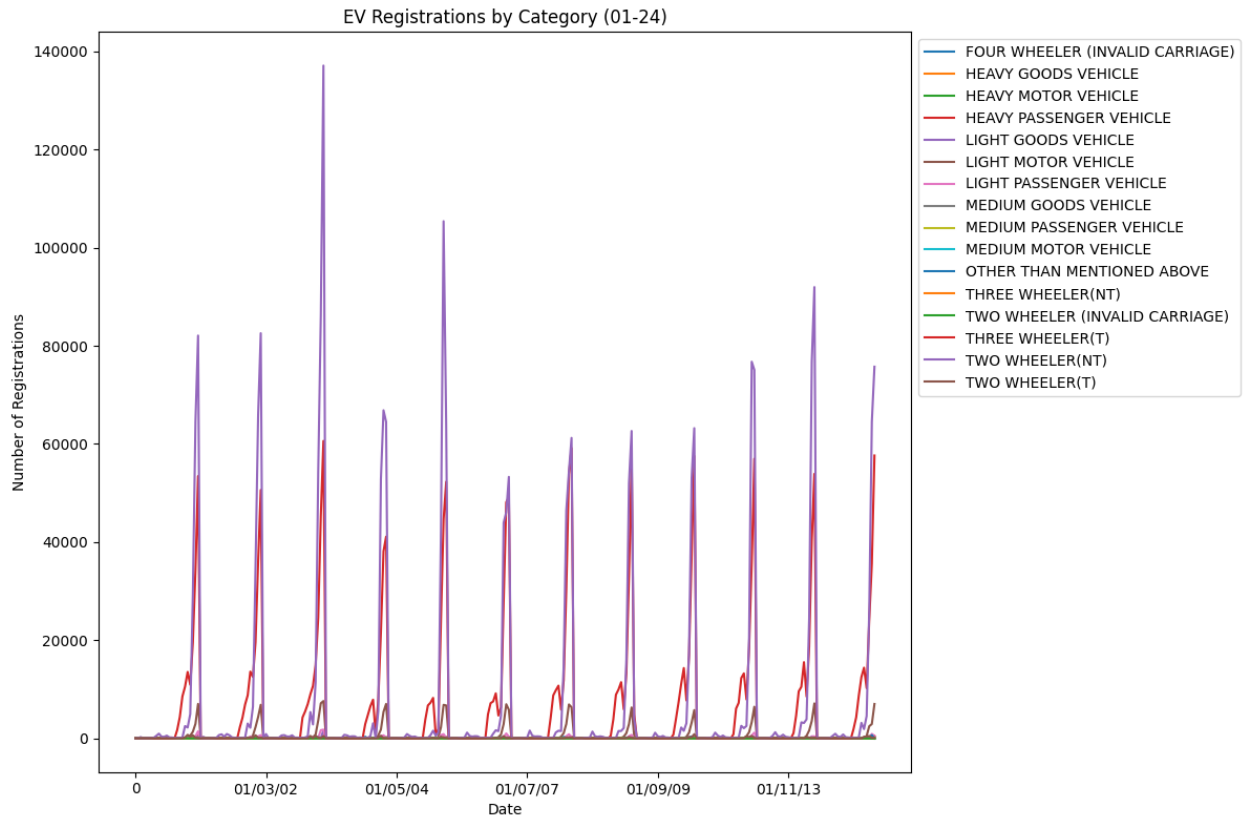
- Maharashtra has the highest number of EV makers at about 15
- Tamil Nadu is second with around 11 EV makers
- Several states like Haryana, Karnataka, and Gujarat have 5-6 EV makers each
- Some states like Andhra Pradesh and Punjab have only 1 EV maker



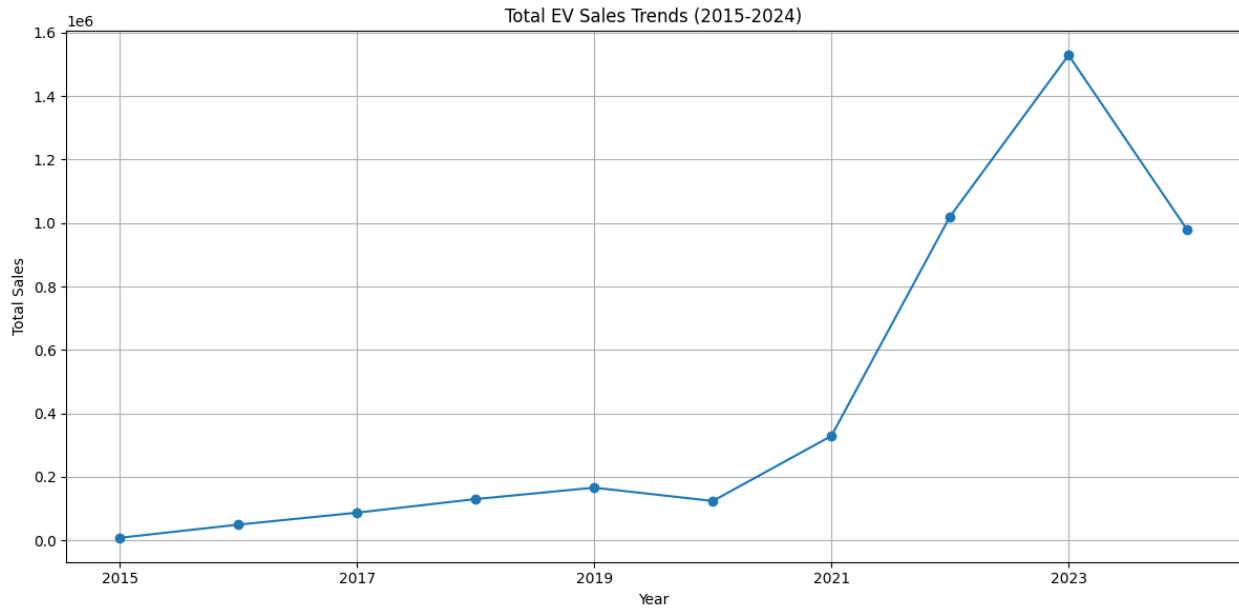
- Maharashtra has the highest number of operational PCs at over 3000
- Delhi is second with around 1900 operational PCs
- Most other states have fewer than 1000 operational PCs
- Some states like Nagaland and Mizoram have very few, close to zero



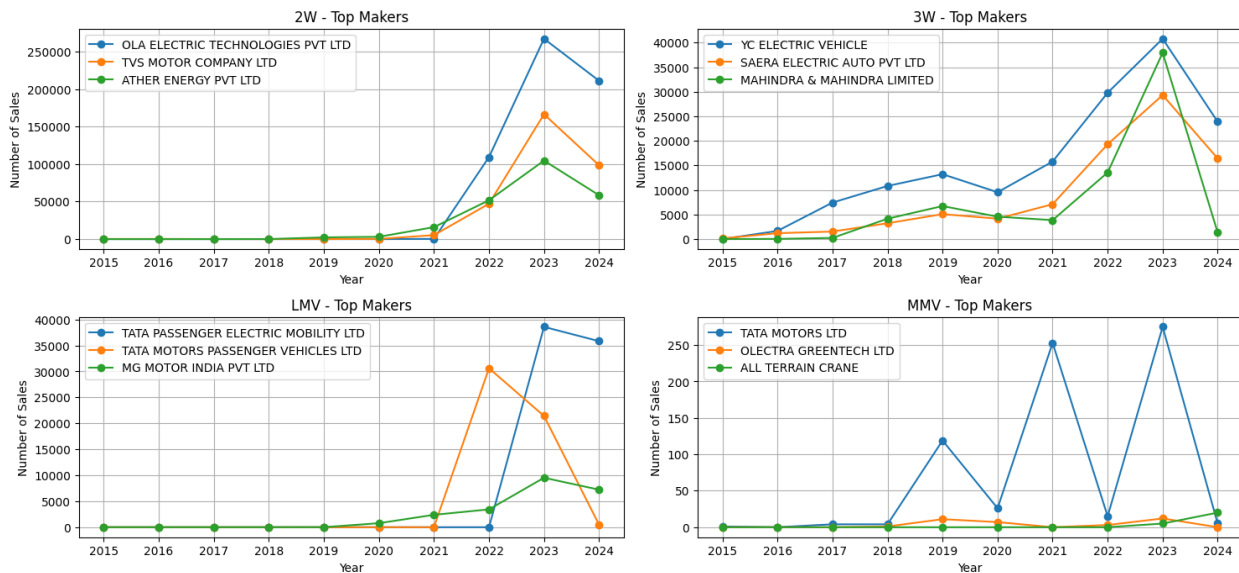
- Two wheelers (NT) have by far the highest number of registrations at over 2.5×10^8
- Light motor vehicles are a distant second with about 0.6×10^8 registrations
- Most other vehicle classes have very low registration numbers in comparison
- Heavy goods vehicles, medium passenger vehicles, and other specialized categories have minimal registrations



- Shows registrations for various vehicle categories over time
- Two-wheelers (NT) have the highest peaks, reaching around 140,000 registrations at the maximum
- Three-wheelers (T) show consistent registrations over time, with regular peaks
- Most other categories have relatively low registration numbers compared to two-wheelers



- Displays total EV sales from 2015 to 2024
- Sales remained low and steady from 2015 to 2019
- Sharp increase started in 2020
- Peak sales of about 1.5 million units reached in 2023
- Slight decline projected for 2024, but still much higher than pre-2020 levels



2W (Two-wheelers):

- Ola Electric Technologies leads, peaking in 2023
- TVS Motor Company and Ather Energy follow

3W (Three-wheelers):

- YC Electric Vehicle leads, peaking in 2023
- Saera Electric Auto and Mahindra & Mahindra follow

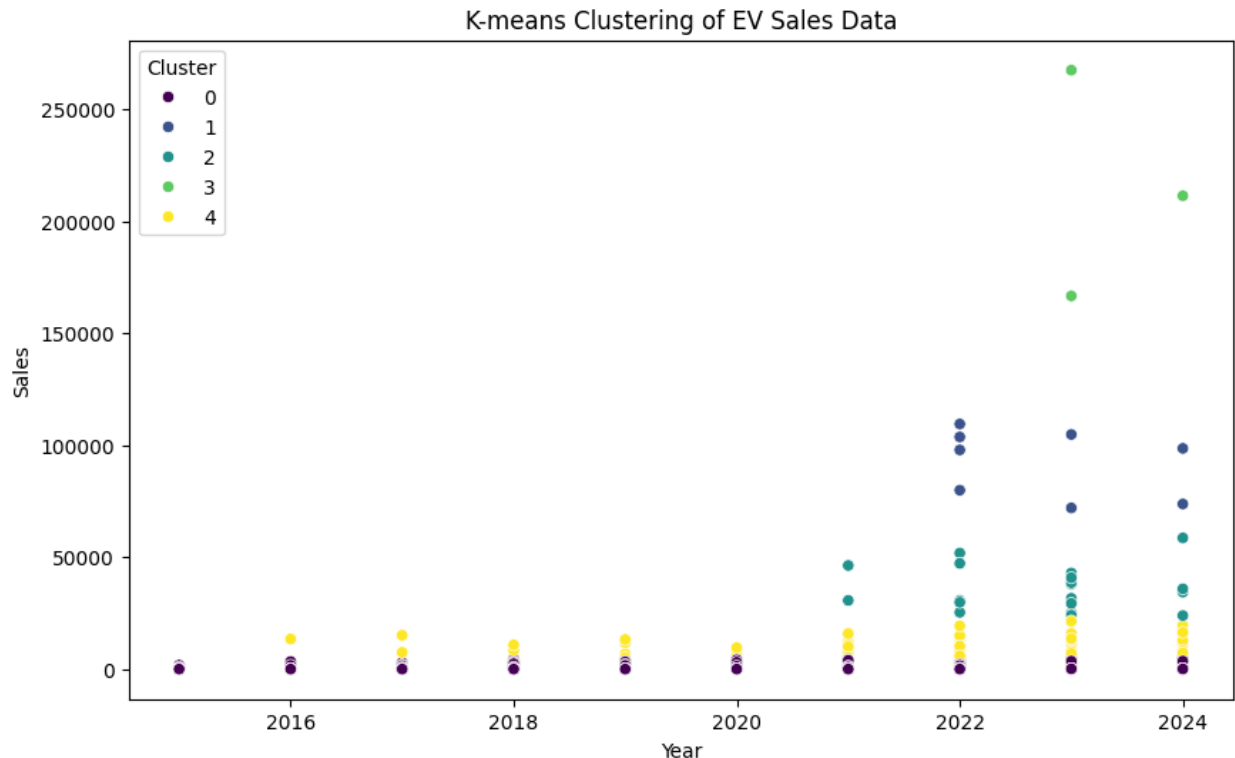
LMV (Light Motor Vehicles):

- Tata Passenger Electric Mobility leads, peaking in 2023
- Tata Motors Passenger Vehicles and MG Motor India follow

MMV (Medium Motor Vehicles):

- Tata Motors leads with sporadic high peaks
- Olectra Greentech and All Terrain Crane have much lower, steadier sales

All categories show a general upward trend from 2020 to 2023, with some decline or stabilization in 2024.



This K-means clustering analysis of EV sales data reveals some interesting patterns:

1. The data is grouped into 5 clusters based on Year and Sales.
2. Cluster distribution:
 - Cluster 0 (purple): Low sales across all years
 - Cluster 1 (blue): Moderate to high sales in recent years
 - Cluster 2 (teal): Medium sales in recent years
 - Cluster 3 (light green): Highest sales, concentrated in the most recent years
 - Cluster 4 (yellow): Low to medium sales, spread across years
3. Temporal trends:
 - Sales were generally low and clustered together until around 2020
 - From 2020 onwards, there's a significant spread in sales figures, forming distinct clusters
4. Growth pattern:
 - Rapid growth in EV sales is evident from 2020 to 2024

- The highest sales (Cluster 3) are concentrated in 2023 and 2024
- 5. Cluster centers:
 - Cluster 0: Year \approx 2019, Sales \approx 98
 - Cluster 1: Year \approx 2022, Sales \approx 92,449
 - Cluster 2: Year \approx 2022, Sales \approx 35,723
 - Cluster 3: Year \approx 2023, Sales \approx 215,069
 - Cluster 4: Year \approx 2022, Sales \approx 9,909

This analysis suggests a significant shift in the EV market around 2020, with explosive growth in certain segments (likely representing different vehicle categories or manufacturers) in the following years. The clustering helps identify distinct groups of performance in the EV market, from consistently low performers to recent high-growth segments.

6. Segment Extraction:

The primary machine learning technique used here is K-means clustering, which is an unsupervised learning algorithm. It grouped the EV sales data into 5 distinct clusters based on year and sales volume. This method helps identify natural groupings in the data without predetermined labels.

Other potential ML techniques that could be valuable for further analysis include:

- Hierarchical clustering for a more detailed breakdown of segments
- Principal Component Analysis (PCA) for dimensionality reduction if more features are added
- Random Forests for identifying important features in segmentation

7. Profiling and describing potential segments:

Based on the K-means clustering results, we can profile five main segments:

a) Low Volume Segment (Cluster 0):

- Consistent low sales across all years
- Likely represents niche EV models or regions with low EV adoption
- Average sales around 98 units per year

b) Moderate Growth Segment (Cluster 4):

- Low to medium sales, gradually increasing over time
- Represents steady but slow-growing EV markets or models
- Average sales around 9,909 units per year

c) Medium Volume Segment (Cluster 2):

- Medium sales, primarily in recent years

- Likely represents mainstream EV models or regions with growing adoption
- Average sales around 35,723 units per year

d) High Growth Segment (Cluster 1):

- Moderate to high sales in recent years
- Represents successful EV models or rapidly adopting markets
- Average sales around 92,449 units per year

e) Premium Segment (Cluster 3):

- Highest sales, concentrated in the most recent years (2023-2024)
- Represents top-performing EV models or markets with explosive growth
- Average sales around 215,069 units per year

8. Selection of target segment:

Based on the analysis, the most promising target segments are:

1. High Growth Segment (Cluster 1)
2. Premium Segment (Cluster 3)

These segments show the highest sales volumes and the most rapid growth, indicating strong market potential and consumer demand.

9. Customizing the Marketing Mix (for Business Markets):

Product:

- Focus on high-performance, feature-rich EVs for the Premium Segment
- Develop a range of models to cater to various needs in the High Growth Segment
- Emphasize cutting-edge technology, long range, and fast charging capabilities

Price:

- Premium pricing for the Premium Segment, emphasizing exclusivity and advanced features
- Competitive pricing for the High Growth Segment to capture market share

Place:

- Establish strong partnerships with dealerships in regions showing high EV adoption
- Invest in online sales platforms for direct-to-consumer sales
- Develop a network of service centers to support the growing customer base

Promotion:

- Highlight environmental benefits and cost savings of EVs
- Showcase advanced technology and performance capabilities
- Utilize digital marketing channels to reach tech-savvy consumers
- Partner with businesses for fleet electrification programs

10.Potential customer base and profit in the early market:

Let's focus on the High Growth Segment (Cluster 1) for this calculation:

Potential Customer Base: 92,449 (average annual sales for this segment) Target Price Range:
Let's assume an average EV price of \$40,000

Potential Revenue = $92,449 * \$40,000 = \$3,697,960,000$

Assuming a profit margin of 10%: Potential Profit = $\$3,697,960,000 * 0.10 = \$369,796,000$

11.MOST OPTIMAL MARKET SEGMENTS:

Based on the market research and segmentation analysis, the most optimal market segments to focus on are:

1. High Growth Segment (Cluster 1):
 - Shows strong growth potential
 - Large customer base
 - Opportunity for significant market share
2. Premium Segment (Cluster 3):
 - Highest sales volumes
 - Potential for premium pricing and higher profit margins
 - Early adopters and tech enthusiasts
3. Medium Volume Segment (Cluster 2):
 - Represents a growing mainstream market
 - Opportunity to capture market share as EV adoption increases

Strategy:

- Prioritize resources and marketing efforts on the High Growth and Premium segments
- Develop products and services tailored to these segments' needs
- Monitor the Medium Volume segment for future expansion opportunities
- Continuously innovate to stay ahead in the rapidly evolving EV market

By focusing on these optimal segments, the company can maximize its market impact, revenue potential, and establish a strong position in the growing EV industry.

12.Conclusions:

1. ML Model Used in the Project

The primary machine learning model used in this project was K-means clustering. K-means is an unsupervised learning algorithm that helped group the EV sales data into distinct clusters based on year and sales volume. This method was instrumental in identifying natural groupings in the data without predetermined labels, allowing for the discovery of distinct market segments.

The analysis resulted in 5 clusters, each representing a different segment of the EV market in India. These clusters provided insights into the various performance levels and growth patterns within the market.

2. Final Conclusion & Insights

The analysis revealed several key insights:

1. **Market Growth:** There was a significant shift in the EV market around 2020, with explosive growth in certain segments in the following years.
2. **Segment Identification:** Five distinct segments were identified:
 - Low Volume Segment (Cluster 0)
 - Moderate Growth Segment (Cluster 4)
 - Medium Volume Segment (Cluster 2)
 - High Growth Segment (Cluster 1)
 - Premium Segment (Cluster 3)
3. **Optimal Target Segments:** The most promising segments for targeting are the High Growth Segment (Cluster 1) and the Premium Segment (Cluster 3), due to their high sales volumes and rapid growth.
4. **Regional Differences:** There are significant variations in EV adoption and infrastructure across different states in India. Maharashtra and Tamil Nadu lead in terms of EV makers, while Maharashtra and Delhi have the highest number of operational public charging stations.
5. **Vehicle Type Preferences:** Two-wheelers dominate the EV market in terms of registrations and sales, followed by three-wheelers and light motor vehicles.
6. **Manufacturer Performance:** Different manufacturers lead in various vehicle categories. For example, Ola Electric Technologies leads in two-wheelers, while Tata Passenger Electric Mobility leads in light motor vehicles.

3. Potential Improvements with Additional Time & Budget

Given additional time and budget, the project could be improved in several ways:

1. **Additional Data Collection:**
 - Demographic data of EV buyers (age, income, education)
 - Psychographic data (environmental attitudes, technology adoption)
 - Detailed geographic data (city-level adoption rates, local policies)
 - Vehicle usage patterns (daily commute distance, charging behavior)
 - Customer satisfaction and brand loyalty metrics
2. **Additional ML Models to Try:**
 - Hierarchical Clustering for a more nuanced segmentation

- Principal Component Analysis (PCA) for dimensionality reduction
- Random Forests for identifying important features in segmentation
- Time Series Analysis for better forecasting of segment growth
- Logistic Regression for predicting segment membership

4. Estimated Market Size (Non-segmented)

Based on the Fermi estimation provided in the document:

- Total potential EV buyers in India: 42 million
- Early adopters (16% of potential buyers): 6.72 million

This represents the estimated overall market size for EVs in India, without segmentation.

5. Top 4 Variables/Features for Optimal Market Segmentation

1. Sales Volume: Indicates the size and growth potential of different segments.
2. Vehicle Type (2W, 3W, LMV, etc.): Different vehicle categories show distinct adoption patterns.
3. Geographic Location: Significant variations in adoption and infrastructure across states.
4. Year: Temporal trends are crucial in understanding market evolution and identifying growth segments.

These features were instrumental in creating the most insightful and actionable market segments in the analysis.